The Tube over Time: Characterizing Popularity Growth of YouTube Videos

Flavio Figueiredo   Fabrício Benevenuto   Jussara M. Almeida
{flaviov, fabricio, jussara}@dcc.ufmg.br
Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte - Brazil

ABSTRACT

Understanding content popularity growth is of great importance to Internet service providers, content creators and online marketers. In this work, we characterize the growth patterns of video popularity on the currently most popular video sharing application, namely YouTube. Using newly provided data by the application, we analyze how the popularity of individual videos evolves since the video's upload time. Moreover, addressing a key aspect that has been mostly overlooked by previous work, we characterize the types of the referrers that most often attracted users to each video, aiming at shedding some light into the mechanisms (e.g., searching or external linking) that often drive users towards a video, and thus contribute to popularity growth. Our analyses are performed separately for three video datasets, namely, videos that appear in the YouTube top lists, videos removed from the system due to copyright violation, and videos selected according to random queries submitted to YouTube's search engine. Our results show that popularity growth patterns depend on the video dataset. In particular, copyright protected videos tend to get most of their views much earlier in their lifetimes, often exhibiting a popularity growth characterized by a viral epidemic-like propagation process. In contrast, videos in the top lists tend to experience sudden significant bursts of popularity. We also show that not only search but also other YouTube internal mechanisms play important roles to attract users to videos in all three datasets.

Categories and Subject Descriptors

C.4 [**Computer Systems Organization**]: Performance of Systems—*Measurement techniques*; H.3.5 [**Information Storage and Retrieval**]: Online Information Services—*Web-based services*

General Terms

Human Factors, Measurement

Keywords

YouTube, video popularity, popularity growth, referrer

1. INTRODUCTION

Understanding content popularity growth on the Internet is of great relevance to a broad range of services, from technological, economical and social perspectives. Such understanding can drive the design of cost-effective caching and content distribution mechanisms as well as uncover potential bottlenecks in system components such as search engines [6]. Moreover, predicting popularity is also important not only for supporting online and viral marketing strategies as well as effective information services (e.g., content recommendation and searching services) [12] but also because it may uncover new (online and offline) business opportunities. From a sociological point of view, a deep study of popularity evolution may also reveal properties and rules governing collective user behavior [10].

Online Social Networks (OSNs) are currently a major segment of the Internet. Considering video sharing OSNs, YouTube[1] is the one with the largest number of registered users [1], who upload and share their videos at a staggering rate. Indeed, it has been reported that the amount of content uploaded to YouTube in 60 days is equivalent to the content that would have been broadcasted for 60 years, without interruption, by NBC, CBS and ABC altogether [2]. Moreover, YouTube has reportedly served over 100 million users only on January 2009 [1], with a video upload rate equivalent to 10 hours per minute [2]. At such unprecedented user and content growth rates, understanding video popularity on YouTube becomes a challenge of utmost importance, as the myriad of different contents make user behavior and attention span highly variable and unpredictable [6].

As argued by Willinger *et al.* [20], most previous analyses of OSNs have treated such systems as static. Most of them focus on analyzing structural properties of single snapshots of relationship networks (e.g., friendship network) that emerge in such systems [3, 5, 15]. However, since OSNs are inherently dynamic, these studies fail to address key properties of the underlying system dynamics. Regarding one such property, namely popularity, a few studies have analyzed YouTube with respect to video popularity characteristics [6, 9, 10] and prediction [14, 19]. However, most of them, despite covering a rich set of popularity properties and their implications for system design, focused on only a single or

---

[1] http://www.youtube.com
[2] http://www.youtube.com/t/fact_sheet